

## **METHOD AND SYSTEM FOR MULTIPLE PASS VIDEO CODING**

### **BACKGROUND OF THE INVENTION**

#### **Field of the Invention**

[0001] The present invention relates generally to the field of data compression, and more specifically to a method and system for real-time multiple-pass data encoding, and particularly, for encoding video signals.

#### **Description of the Prior Art**

[0002] Digital video compression is an essential technology in video communications, broadcasting, and storage. MPEG video coding standards have been successfully used to reduce the transmission bandwidth and storage space requirements in many applications, such as digital TV broadcast through satellite and cable, Digital Video Disk (DVD), Video on-Demand and video streaming over the Internet, etc. However, emerging applications and new services become increasingly demanding for less transmission bandwidth and storage space. For example, in Video-on-Demand service over Asymmetric Digital Subscriber Line (ADSL), live news and sports events are transmitted in real-time to the subscribers using MPEG-2 video coding standard (ISO/IEC 13818-2) at a constant bit rate (CBR) in the range of 0.6 to 2Mbits/second. For MPEG CBR video encoding at such a bit rate range, it is very challenging for the conventional MPEG encoders available on the commercial market to produce acceptable picture quality. Conventional MPEG encoders employ a single encoder scheme as depicted in Figure 1. As shown in Figure 1, the conventional video encoder device 110 implements a coding strategy which is based on the information retained in coding only the previous video frames 100 to provide coded video output 120. These encoders 110 routinely adopt a coding strategy that is only based on the information obtained in coding of the previous video frames 100 and/or rely on some assumed signal models to predict or estimate the signal properties of the current input frame to guide the encoding process of the current frame. However, natural video is a statistically non-stationary signal source.

Prediction and estimation based on the past signal will not correctly describe the current input signal. In addition there is no known robust signal model that can describe the natural video signal reliably. Such encoders will not be able to determine and apply the best coding strategy to encode the incoming video frames for lack of the information about the current and future input frames. In order to meet the challenges from more demanding emerging applications, more sophisticated schemes for MPEG-2 video coding are needed to improve the performance and to ensure the quality of services.

[0003] Research efforts have been made to improve the variable bit rate (VBR) MPEG video coding, e.g., for DVD applications, by employing two-pass and re-encoding schemes. However, there are no published research results for multi-pass CBR coding in the literature.

[0004] It would thus be highly desirable to provide a real-time MPEG CBR video coding method and associated system that are able to jointly determine and apply the best coding plan to encode input video frames based not only on the information of the previous and current frames, but also the information about the future input frames.

#### Summary of the Invention

[0005] According to one aspect of the invention, a system and method is provided for performing real-time multi-pass data encoding, in particular video signal multi-pass encoding.

[0006] According to another aspect of the invention, a system and method is provided for performing real-time video signal multi-pass encoding with information look-ahead.

[0007] According to a further aspect of the invention, there is provided a system and method for MPEG video coding with information look-ahead that utilizes two MPEG encoder devices. The first encoder device functions as an information collector, which information is then used by an on-line processor. Taking the advantage of the time delay between the inputs of the two encoder devices, the processor employs an efficient algorithm to jointly derive the best coding strategy for the all incoming frames in a look-ahead window by exploiting the information not only about the past and current frames but additionally the

future frames. The second encoder, which operates at the same constant bit rate as the first encoder, uses the coding strategy from the processor as the guide to encode the incoming frames.

[0008] Advantageously, the system and method of the present invention may be applicable for encoding any type of digital information that can be divided into coding units having bits that may be allocated to the coding units for constant bit rate or variable bit rate encoding. For example, digital audio or digitized speech can be divided into frames in millisecond units. These frames can be treated the same as the video pictures and the invention can be applied to these coding units.

### Brief Description of the Drawings

[0009] The objects, features and advantages of the present invention will become apparent to one skilled in the art, in view of the following detailed description taken in combination with the attached drawings, in which:

[0010] Figure 1 is a description of a conventional simple encoder system;

[0011] Figure 2 illustrates an encoder system that has look-ahead information collection; and,

[0012] Figure 3 is a block diagram depicting a preferred embodiment for encoding digital data according to the present invention.

### Detailed Description of the Preferred Embodiments

[0013] As illustrated in Figure 2, an information look-ahead mechanism is added to the conventional video signal encoder device of Figure 1. That is, as shown in Figure 2, video input frames 200 are fed in parallel to a buffer device 210 in front of the encoder device 220 and, an information collector/processor device 230. The buffer device 210 functions to delay the input video frames 200 by a fixed amount of time so that the information collector/processor 230 will have the operation time to extract useful information about the

incoming frames in the delay buffer 210 and process the information to determine a coding strategy for encoding these frames. The determined coding strategy is then passed in the form of coding parameters to the encoder device 220 for execution.

**[0014]** In the video coding scheme depicted in Figure 2, the buffer effectively creates a look-ahead time window for the information collector/processor to gather and process the information. Given the information and the processing algorithm executed in the information collector/processor device 230, the best coding plan may be determined jointly based on the information about the past, current, and future input video frames. However, to implement this video coding scheme in a cost effective way and to achieve efficient performance, it is necessary to first determine what appropriate look-ahead information to collect from the input frames and how to collect it, how the look-ahead information should or can be used to devise the best coding strategy (i.e., the processing algorithm) and how the coding strategy can be carried out, and what is the proper buffer size (or the look-ahead window size).

**[0015]** The most useful information for determining the best coding strategy for the incoming video frames are the signal statistics and characteristic variables, rate-quality measure, and coding parameters that are directly used in various steps of the encoding process with the dominant impact in the coding results. The most effective approach to collect such information is to use a collector that emulates the encoder operation. Therefore, in order to gather the most pertinent and useful information to derive the best coding strategy, a second MPEG encoder device is employed such as depicted in the block diagram of Figure 3 illustrating the preferred system of the invention. In the system depicted in Figure 3, two MPEG encoder devices 320 and 330 are provided that operate at the same CBR. Particularly, video input frames 300 are fed in parallel to a buffer device 310 and, a first MPEG encoder device 330 which functions as the information collector and feeds the processor device 340 which implements a processing algorithm executed within the information processor 340 which may be implemented in a general processor, a DSP chip, or reside on a host PC (not shown). The primary benefit of using the first MPEG encoder 330 as the information collector is that the direct signal information and intermediate results in various encoding stages can be obtained in the same encoding operation conditions as the intended encoding process. The exact items of information to be collected may depend on the need of the

processing algorithm and the availability of the information in the encoder chip on the fly and the real-time output capability of the encoder device 330. In the preferred embodiment depicted in Figure 3, information relating to the picture complexity, motion magnitude, and picture quality index are collected.

[0016] As an initial consideration, the length of the look-ahead window determines the input delay buffer 310 size. The more frames to look-ahead, the larger the buffer size and in turn the longer the delay. The cost also increases with the buffer size. For the convenience of bit allocation and rate control in CBR coding, the look-ahead window size is a predetermined multiple of the size of Group of Pictures (GOP) so that the numbers of Intra-coded (I), Predictive Coded (P), and Bi-direction Predictive Coded (B) frames in the look-ahead window are constants. Details regarding the MPEG encoding frames may be found in a reference entitled *Informaiton Technology—Generic Coding of Moving Pictures and Associated Audio:Video*, ISO/IEC 13818-2, 1995 incorporated by reference as if fully set forth herein. The look-ahead window size  $W_s$  is thus determined to be:

$$[0017] \quad W_s = K * GOP_s$$

where  $K = 1$  or  $2$ ;  $GOP_s$  is the size of Group of Picture in MPEG video coding. The input delay buffer size  $B_s$ , then becomes:

$$[0018] \quad B_s = W_s + \Delta_p$$

where  $\Delta_p$  is the information processing time which depends on the complexity of the algorithm.

[0019] Once the information about the video frames in the look-ahead window is available, the processing algorithm determines a coding strategy for these frames using the information. In the preferred embodiment, a target bit allocation plan for the video frames is jointly determined so that the available bits can be used efficiently and the decoding buffer defined as Virtual Buffer Verifier (VBV) in MPEG-2 standard can be exploited sufficiently. Assume there are  $N$  frames in the look-ahead window. Let  $P_i$ ,  $i=1, \dots, N$ , be the  $i$ -th frame in the

window. The picture complexity, motion magnitude, picture quality index, and target number of bits for  $P_i$  are denoted as  $C_i$ ,  $M_i$ ,  $Q_i$  and  $T_i$ , respectively. With  $R$  representing the bit rate and  $F$  the frame rate, the algorithm performs the following steps:

[0020] A first step is to calculate the dynamic weighted picture complexity,  $C_i$  as:

$$[0021] \quad C'_i = C_i W\left(\frac{M_i}{\bar{M}_I}, \frac{Q_i}{\bar{Q}_I}, S_i, D_i\right);$$

where  $W()$  is a real function;  $S_i \in \{I, P, B\}$  is the picture coding type of frame  $P_i$ ;  $\bar{M}_I$  and  $\bar{Q}_I$  are the average motion magnitude and average picture quality index of all frames in the look-ahead window with the same picture coding type as  $S_i$ ; and  $D_i$  is the distance from  $P_i$  to the most recent I frame. It should be noted that the larger the value of  $Q_i$ , the worse the picture quality.

[0022] A second step is to jointly determine the target number of bits for all frames in the look-ahead window:

$$[0023] \quad T_i = \frac{RNC'_i}{F \sum_{k=1}^N C'_k};$$

[0024] A third step is the step of determining rate control to prevent decoder buffer overflow and underflow. The variable “V” is denoted as the decoder buffer size (e.g., 1835008 bits for MP@ML (Main Profile/Main level) case) as defined in MPEG-2 standard (See *Informaiton Technology—Generic Coding of Moving Pictures and Associated Audio:Video*, ISO/IEC 13818-2, 1995) and  $V_i$  is denoted as the decoder buffer fullness just before the picture  $P_i$  is drawn from the decoder buffer for decoding. Letting “G” be a guard band, for example,  $G = 3\% \sim 5\%$  of  $V$ , the MPEG-2 decoder buffer model for CBR operation is described by the following recurrence:

$$[0025] \quad \begin{aligned} V_0 &= V_{init}, \\ V_i &= V_{i-1} + R/F - T_i, \end{aligned}$$

where  $V_{\text{init}}$  is the initial buffer fullness. To prevent any overflow and underflow, the buffer fullness must always satisfy the following relation:

$$[0026] \quad T_i + G \leq V_i \leq V - G.$$

[0027] If  $V_i$  underflow or overflow the buffer requirement by an amount of  $\delta$ , then the target bit allocation must be adjusted according to the following:

$$[0028] \quad \begin{aligned} T'_k &= T_k \frac{\sum_{m=1}^i T_m \mp \delta}{\sum_{m=1}^i T_m}, & k &= 1, 2, \dots, i; \\ T'_k &= T_k \frac{\sum_{n=i+1}^N T_n \pm \delta}{\sum_{n=i+1}^N T_n}, & k &= i+1, \dots, N. \end{aligned}$$

[0029] Returning to Figure 3, after the coding strategy is determined for the video frames in the look-ahead window, it needs to be passed to the second encoder device 320. This is preferably communicated in the form of the coding parameters and other necessary information. In practice, the real-time communication bandwidth between the second encoder device 320 and information processor 340 may limit the amount of the coding parameters and information to be transmitted to the encoder 320 on frame-by-frame basis and may have an impact on the execution of the coding strategy. In one of the embodiments of the invention, four (4) 16 bit integers are used to pass parameters to the second encoder 320 for every frame's encoding. They are 16 bits for the target number of bits, 16 bits for the weighted picture complexity, and 32 bits for the sum of the weighted complexities of the remaining un-encoded frames in the look-ahead window. The last two parameters are used by the second encoder 320 to reallocate any excess bits over the remaining frames.

[0030] While the invention has been described for MPEG video encoding, it is understood that the invention may be used with other video coding techniques or even with non-video data. Indeed, any digital information can be divided into coding units and bits are allocated to the coding units for constant bit rate or variable bit rate encoding. For example, digital audio or digitized speech can be divided into frames in millisecond units. These frames can be treated the same as the video pictures and the invention can be applied to these coding units.

[0031] While the invention has been particularly shown and described with respect to illustrative and preformed embodiments thereof, it will be understood by those skilled in the art that the foregoing and other changes in form and details may be made therein without departing from the spirit and scope of the invention which should be limited only by the scope of the appended claims.